Age estimation for Asian Cymbidium (Orchidaceae: Epidendroideae) with implementation of fossil data calibration using molecular markers (ITS2 & matK) implying phylogeographic inference

Devendra K Biswal, Jean V Marbaniang, Pramod Tandon

Abstract: Intercontinental dislocations between tropical regions harboring two-thirds of the flowering plants have always drawn attention from taxonomists and bio-geographers. The focus had always been on woody land plants rather than on herbs. Orchidaceae is one such family belonging to angiosperms, with an herbaceous habit and high species diversity in the tropics. Here, we investigate the evolutionary and bio-geographical history of the genus *Cymbidium*, which represents a monophyletic subfamily (Epidendroideae) of the orchids and comprises 50 odd species that are disjunctly distributed in tropical to temperate regions. A relatively well-resolved and highly supported phylogeny of *Cymbidiums* was reconstructed based on sequence analyses of internal transcribed spacer (ITS2) regions and maturaseK (matK) from the chloroplast region available on the public domain in GenBank at NCBI. Crassulacean acid metabolism (CAM) is one of the photosynthetic pathways regarded as adaptation to water stress in land plants. Hardly any information exists on correlations among the level of CAM activity, habitat, life form, and phylogenetic relationship of a plant group from an evolutionary perspective. This study examines a genus level analyses by integrating ITS and matK data to all fossil data available on orchids in a molecular Bayesian relaxed clock employed in BEAST and assessed divergence times for the genus *Cymbidium* with a focus on evolutionary plasticity of photosynthetic characters. Our study has enabled age estimations for the genus *Cymbidum* (12Ma) for the first time using BEAST by addition of previously analyzed two internal calibration points.

PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

Devendra K Biswal[1], Jean Valrie Marbaniang[2], Pramod Tandon[1,2,§]

1Bioinformatics Centre, North-Eastern Hill University, Shillong, Meghalaya, India
2Department of Botany, Centre for Advanced Studies , North-Eastern Hill University, Shillong, Meghalaya, India
§Corresponding author

Email addresses:
DKB: devbioinfo@gmail.com
JVM: jean_valrie@yahoo.co.in
PT: profptandon@gmail.com

§Corresponding author (s)

Prof. Pramod Tandon
Address:
Department of Botany, Centre for Advanced Studies
North-Eastern Hill University
Shillong 793022, India
Tel. Work +913642550150
+913642722244
Fax. +913642550300
E-mail: profptandon@gmail.com

Devendra Kumar Biswal
Address:
Bioinformatics Centre
North-Eastern Hill University
Shillong 793022, India
Tel. Work +91364 2721241
Fax. +91364 2550300
E-mail: devbioinfo@gmail.com, devendra@nehu.ac.in

PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

Introduction:

Intercontinental dislocations between tropical regions harboring two-thirds of the flowering plants have always drawn attention from taxonomists and biogeographers. The focus had always been on woody land plants rather than on herbs. Orchidaceae is one such family belonging to angiosperms, with a herbaceous habit and a high species diversity in the tropics. *Cymbidium* Sw. is a popular group of orchids including some 50 species and possess significant commercial importance in the floriculture industry. Their flowers can come in all shades of colors with varying attractive shapes of high durability and large size (Du Puy and Cribb, 2007). The species distribtuion of this genus range from Northeast India to China and Japan, south through the Malay Peninsula to Northern and Eastern Australia where various types of vegetation such as tropical to cool–temperate rainforest, tropical to subtropical seasonal forest, savanna woodland, subtropical grassland, and tropical to subtropical montane forest are established. The larger flowered species grow at high altitudes (Du Puy and Cribb, 2007). Altitudes range from 0 to 3,000 m and latitudes from 44°N to 32°S. Such a wide range of space offers various combinations of environmental conditions such as water, light, temperature, CO_2 and nutrients. In addition to this three different life forms (terrestrial, epiphytic, and lithophytic) and two feeding properties (autotrophic and mycoheterotrophic) have developed within the genus under these environmental regimes. These facts indicate that ecological and physiological characteristics are also diverse within this genus.

Cymbidium orchids were among the earliest to be cultivated, especially in China. Though they are not widely cultivated, *Cymbidium* hybrids lend themselves to cultivation; commercially important hybrids have been produced for over hundred years and make excellent ornamental and pot plants for which they have been the subject of taxonomic studies and species identification. Traditionally, the species taxonomy and cultivars identification of the genus *Cymbidium* was entirely based on morphological traits which is very difficult, as morphology is affected by environmental factors and as of yet, there are no efficient methods for identifying *Cymbidium* species and cultivars. Consequently, species identification and study of *Cymbidium* evolution and taxonomy remains a challenge (van den Berg *et al*., 2002; Sharma *et al*., 2012).

Molecular methods, such as molecular marker techniques, molecular morphometrics and phylogenetics and DNA barcoding, provide effective information for taxonomy, species identification and phylogenetics. Choi *et al.* (2006) used RAPD markers for identification of *Cymbidium* cultivars and their relationships. Wang *et al.* (2011) reported the cultivar identification in *C. ensifolium* using ISSR markers. van den Berg *et al.* (2002) used ITS and matK to elaborate the phylogenetic relationships of *Cymbidium* while Sharma *et al*. (2012) assessed the phylogenetic inter-relationships of *Cymbidium* taking advantage of ITS. Most of these studies agreed upon the fact that a limited number of DNA sequences led to relatively little genetic variation within the genus and therefore phylogenetic resolution and species identification was very difficult.

Photosynthetic characters for *Cymbidium* species have been identified (Wong and Hew, 1973; Motomura *et al*., 2008) and these data suggest that the genus comprises both C3 and CAM species. A fully revised *Cymbidium* taxonomy (Du Puy and Cribb, 1988) and their phylogenetic relationships in the genus and allied genera have been clarified on the basis of macromolecular characters. Correlations between anatomical and ecological characters have also been discussed(van den Berg *et al*., 2002). Consequently, *Cymbidium* provides a unique opportunity to understand the adaptive and evolutionary significance of photosynthetic characters from an evolutionary point of view. Orchidaceae that happens to be the largest and one of the ecologically and morphologically most diverse families of flowering plants have been assigned to the origin of modern orchid lineages (i.e., their crown age), ranging from ~26 million years (Ma) (Wikström *et al*., 2001), ~40 Ma (Bremer, 2000), ~80 Ma (Ramirez *et* *al*., 2007) to as much as ~110 Ma (Janssen and Bremer, 2004). A correct time estimation is essential for our understanding of the mechanisms underlying the diversification of orchids, and *Cymbidium* could significantly contribute towards evolutionary studies owing to its genetic diversity - such as significant increases in speciation rates that may temporally correlate to climatic changes, tectonic events, or radiation of pollinators etc.

Molecular dating of the Orchidaceae has been really challenging due the complete absence of reliable orchid fossils and a recent study by Ramirez *et al*., (2007) hints at a 15-20 Ma fossil of an extinct stingless bee (*Proplebeia dominicana*), covered with pollinia from an orchid species belonging to the subtribe Goodyerinae, that helped temporal calibration of the family. The fossil is used as an internal calibration point, and departing from a phylogenetic tree obtained from the analysis of plastid DNA sequences (matK and rbcL), Ramirez *et al*. (2007) opined on the origin of Orchidaceae at an estimated 76-84 Ma suggesting 'ancient' origin of orchids in the Late Cretaceous. Besides, fossil leaves of two Early Miocene orchids (*Dendrobium* and *Earina*) reported from New Zealand represent the first Orchidaceae macrofossils with cuticular preservation, the oldest records for subfamily Epidendroideae, as well as the first New Zealand and southern hemisphere fossil records for Orchidaceae (Conran *et al*., 2009).

Till date there have been no studies on the age estimation of Asian and South East Asian orchids. The use of DNA sequences to estimate the timing of evolutionary events is increasingly becoming popular. The idea of dating evolutionary divergences using calibrated sequence differences was first postulated in 1965 by Zuckerkandl and Pauling (1965) highlighting the fact that the amount of difference between the DNA molecules of two species is a function of the time since their evolutionary separation. Several dating techniques based on molecular sequence data have been employed across different taxonomic levels, e.g. to infer the age of plastid-containing eucaryotes (Yoon *et al*., 2004), land plants (Sanderson *et* *al*., 2004), tracheophytes (Soltis *et al*., 2002) angiosperms, (Sanderson and Doyle, 2001; Chaw *et al*., 2001) monocot–dicot divergence (Magallon and Sanderson, 2001), Asterids (Bremer *et al*., 2004), Myrtales (Sytsma *et al*., 2004) and Orchidaceae (Gustafsson *et al*., 2010). This encouraged us to carry out a research study on *Cymbidium* as a case study employing *in silico* open source molecular dating tools on all the available data on the public domain to report the first age estimates for the oriental *Cymbidium* with primarily two objectives:

1. Assessing the temporal origin and diversification of the genus *Cymbidium*, using the Bayesian uncorrelated relaxed molecular clock approach implemented in Bayesian evolutionary analysis by sampling trees (BEAST) (Drummond *et al*., 2012) by use of concatenated ITS and matK sequences with two internal calibration points (15-20 Ma fossil data of an extinct stingless bee as a pollinator for the orchid subtribe Goodyerinae and macromolecular leaf 20 Ma fossil data of genera *Dendrobium* and *Earina*).
2. Phylogeographic inference of Asian *Cymbidium* in continuous space with a focus on evolutionary correlation with CAM related activity using Bayesian Evolutionary Analysis Utility (BEAUti) and BEAST (Drummond *et al*., 2012).

Materials and Methods:

Taxon Sampling and Sequence Analysis

The present study comprised two datasets; one 33 matK and 42 ITS (5.8s, ITS1 and ITS2) sequences from *Cymbidium*, *Dendrobium*, *Earina* and *Goodyera* species and one accession from *Gramtophyllum speciosum* retrieved from GenBank (Benson *et al*., 2011) (via Ebot http://www.ncbi.nlm.nih.gov/Class/PowerTools/eutils/ebot/ebot.cgi website, an open source interactive tool that generates a Perl script implementing an E-utility pipeline for retrieving large datasets from National Centre for Biotechnology Information (NCBI) with key words and boolean operators. matK and ITS accessions from *Dendrobium*, *Earina* and *Goodyera* species were retrieved to serve as internal calibration points and *G. speciosum* as an outgroup for divergence rate studies. The second dataset comprised only ITS2 sequences from Asian *Cymbidium* species and two *Cymbidium* species from UK and Florida to serve as outgroups for phylogeography inference in a continuous space using BEAUti and BEAST (Drummond *et al*., 2012). In GenBank, the country field varies in terms of granularity such as some entries have just a country while other entries are much more detailed. The lat_lon field indicates the latitude and longitude coordinate for the location. A GenBank record can include a lat_lon field with or without the country field. In order to identify which records had a sufficient level of geography, all the *Cymbidium* ITS2 sequence entries were retrieved and queries filtered down to just those that had sufficient geographic information viz. location and coordinate values. The results were split into categories based on the contents of the ‘country’ and ‘lat_lon’ fields (having either or both fields). States and provinces in GeoNames are 1st-level administrative boundaries (ADM1), while countries (ADM2), municipalities, and towns have a different administrative code. For this study, we considered any geographical boundary below ADM1 (except for regions) to be sufficient and relied on literature survey for coordinate values. Information on all the species along with GenBank accessions of both the datasets is summarized in Table 1. Sequences were aligned with default option L-INS-I in MAFFT v.6.717b (Katoh and Toh, 2008) and assume that there is one alignable domain that is flanked by difficult to align residues. Each DNA region was aligned individually and poorly aligned regions (characters with more than 50% gaps) were excluded from the analyses with GBLOCKS 0.91b (Catresana, 2000). Congruence Among Distance Matrices (CADM) test was conducted to test the congruence of datasets by using Kendall's W statistic (Leigh *et al*., 2011). We created a smaller dataset that contained only those species available for both the markers because the program can only test congruence of distance matrices of the exact same size. All aligned regions were concatenated into one dataset generating a supermatrix using Mesquite program V2.75 available at http://mesquiteproject.org. Subsequently the concatenated files were converted to nexus formats to be readable by Bayesian methods. See Additional file 1 in supporting information for the super alignment file.

Phylogenetic Analysis-Bayesian Inference

The dataset was analysed with MRBAYES v.3.1.2 (Ronquist and Huelsenbeck, 2003). Two parallel Bayesian analyses with four chains each and partitioned by DNA region were run for 10 million generations, a sample frequency of 1000 and a burn-in of 25%. Bayesian analysis was carried out for tree construction using a general time reversible substitution model (GTR) with substitution rates estimated by MrBayes. Metropolis-Coupled Markov Chain Monte Carlo (MCMCMC) sampling was performed with two incrementally heated chains that were combinatorially run for 20,000 generations. The convergence of MCMCMC was then monitored by examining the value of the marginal likelihood through generations. Coalescence of substitution rate and rate model parameters were also examined. Average standard deviation of split frequencies was checked and the generations were kept on adding until the standard deviation value was below 0.01. Evolutionary models for each DNA region were determined by the Akaike Information Criterion (AIC) with MRAIC v.1.4.3 at http://www.abc.se/~nylander/. TRACER v.1.5 (Rambaut and Drummond, 2007) was used to evaluate mixing of chains and to determine burn-in. Posterior Probabilities (PP) indicate clade support and values of 1.00–0.90, 0.89–0.70 and 0.69–0.50 indicated strong, moderate and weak support respectively. A cladogram with the posterior probabilities for each split and a phylogram with mean branch lengths was generated and subsequently read by FigTree v1.3.1 at http://tree.bio.ed.ac.uk/software/figtree/

PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

Age Estimation, Fossil Data Calibration And Dating Analyses

The previously aligned concatenated supermatrix in nexus format was opened in BEAUti v1.7.3. The Partitions tab lists each partition and their currently assumed substitution model, clock model, and tree. Under separate substitution models, while assuming the clock and tree are linked both the partitions were selected that comprised 2614 and 1893 sites respectively. While both partitions are highlighted the 'Unlink Subst. Models' button were selected and Clock Model and Partition Tree renamed. The Link Clock Models tab was selected to rename the clock models. Similarly, the Partition Tree was renamed. The options in the Taxa window was set to identify internal nodes of interest based on the earlier bayesian tree as a guide. Three monophyletic groups were created viz. *Cymbidium*, *Dendrobium*, *Earina* and *Goodyera* so as to set the calibration nodes. The 'Sites' menu was selected to specify a substitution model for each of our data partitions. Partion 1 was set to GTR, with 'Estimated' base frequencies, and the 'among-site heterogeneity model' to 'Gamma+Invariant sies'. For partition 2 'SRD06' Model was selected as it comprised matK matrix which is a coding gene and thereby partitions the data so that 3rd codon positions are analyzed separately from 1st and 2nd positions and assumes a 'HKY+Γ' model. The Clock Model for our dataset was set to Estimated Lognormal relaxed clock (Uncorrelated). The input data were compiled in BEAUti v1.7.3 with the tree priors set as follows: i) age for the monophyletic subtribe Goodyerinae (corresponding to the age of the fossil orchid pollinia 15 - 20 Ma old;): Exponential prior distribution with a mean of 10 Ma and an offset of 20 Ma; ii) The two additional calibration points for *Dendrobium* and *Earina* set as uniform prior distributions with a lower bound of 20 Ma and an upper bound of 120 Ma (phylogenetic placement following). The age for the root of the tree was set to a normal prior distribution with mean 39 Ma and standard deviation of 5.5 (giving a 95% CI ranging from c. 48 - 49 Ma) corresponding to the resulting age estimate for the 'Higher Epidendroids' based on literature survey. A Yule prior was used to construct the tree and the ucld.mean was adjusted to a uniform prior of 10–0.000001 to reflect reasonable substitution rates per site for plants. In BEAUti, an XML file was generated (Additional file 2) and subsequently two independent runs for each prior setting (uniform and exponential/lognormal) were done in BEAST with 100 million generations. Further details about parameters and prior settings are given in Table 2. All resulting trees were then combined with LogCombiner v1.5.3 (Drummond *et al.*, 2012), with a burn-in of 25%. A maximum credibility tree was then produced using TreeAnnotator v1.5.3 (Drummond *et al.*, 2012), and subsequently visualized.

Phylogeographic Inference In Continuous Space

The second dataset that comprised only the ITS2 sequences for Asian *Cymbidium* with known CAM activity designated as weak CAM, strong CAM and C3 was aligned in clustal w (Thompson *et al*., 1994) and imported into mesquite for generating nexus formatted file that in turn was run in BEAUti v1.7.3. The location data as well as the approximate sampling dates for the sequence data were mined from GenBank records and incorporated into BEAUti using the 'Tips' tab. In the Traits tab, a LatLong.txt tab-delimited file for the *Cymbidium* location data was incorporated into BEAUti v. 1.7.3. The HKY substitution model for the nucleotide data was kept default, base frequencies were set to 'Empirical', and 'use Gamma-distributed rate variation among sites'. The location partition in the Substitution model window was kept with the default Homogenous Brownian model. A Lognormal relaxed molecular clock (Uncorrelated) model was selected and the Rate (initial value) was set to 0.001. Bayesian skyride model was choosen as a flexible demographic tree prior (Coalescent: GMRF Bayesian Skyride) with the default random starting tree. In the Priors tab, the ucld.mean prior was set to a gamma distribution with shape = 0.001 and scale = 1000. In the MCMC tab, the chain length was adjusted to 100,000 and both the sampling frequencies to 100. An additional statistics block (treeDispersionStatistic) was appended in the operators block in the xml file generated by BEAUti that keeps track of the rate of diffusion by measuring the distance covered along each branch (based on the spatial coordinates inferred at the parent and descendent node of each branch), and summing this distance for the complete tree and dividing this by the tree length.

TreeDispersionStatistic block

<treeDispersionStatistic id="dispersionRate" greatCircleDistance="true"> <treeModel idref="treeModel"/> <multivariateTraitLikelihood idref="location.traitLikelihood"/> </treeDispersionStatistic> <!-- END Multivariate diffusion model -->

The BEAUti formatted XML file was run and trees summarized for the longer run using TreeAnnotator (burn-in = 200). The MCC tree was subsequently read in FigTree v1.3.1 and Spatial Phylogenetic Reconstruction of Evolutionary Dynamics (SPREAD-phy1.0.4) (Bielejec *et al.*, 2011). To visualize the oriental *Cymbidium* divergence through time, MCC trees were converted to KML and these trees were projected through time onto a virtual globe using Google Earth at http://earth.google.com.

Results

Phylogenetic Analyses

The CADM test indicated that the individual gene regions are congruent and yielded a Kendal's W statistic of 0.8 (p-value of 0.001). The AIC suggested a GTR model of sequence evolution with GAMMA distribution of rates across sites (GTRG) for ITS, and the addition of invariant sites for matK (GTRIG). Bayes analyses displayed a strongly supported phylogenetic tree with Posterior Probabilities values of 1.00–0.90, 0.89–0.70 and 0.69–0.50 as strong, moderate and weak support respectively (Figure 1). Interestingly all the 36 extant *Cymbidium* species clustered into a crown monophyletic group with strong branch support values that primarily exhibited C3 and weak CAM photosynthetic traits (previously studied from literature survey). Three species namely *C. bicolor* (strong CAM), *C. finlaysonianum* and *C. rectum* clustered into a separate group suggesting CAM plasticity among the various species with an indication of transition from weak CAM to strong CAM enabling the group to advance to tropical lowland thereby enabling them to withstand the harsh and extreme climatic conditions. The stem group included two sub monophyletic groupings of the genera *Goodyera*, *Dendrobium* and *Earina*. The outgroup *Gramatophyllum speciosum* which is a weak CAM supports earlier studies on photosynthetic traits in orchids that suggests weak CAM to be the ancestral state of the genus *Cymbidium*.

Age estimation and Phylogeographic Inference

The maximum credibility tree for the calibrated relaxed molecular clock analysis of the genus *Cymbidium*, using two internal calibration points (15-20 Ma fossil data of an extinct stingless bee as a pollinator for the orchid subtribe Goodyerinae and macromolecular leaf 20 Ma fossil data of genera *Dendrobium* and *Earina*) is shown in Figure 2. Support values for the different clades and branches were high, with only a few values below 0.90 bayesian posterior probability. All *Cymbidium* species clustered together with high support values (posterior probability above 0.90) and supported the monophyly of the genus. However, the intragenic resolution was poor, precisely discerning two subclades. As per our molecular dating, the genus *Cymbidium* fell within the crown group and shared a most recent common ancesstor (MRCA) in the Late Miocene (Tertiary, Cenozoic), ~13.6 Ma (Figure 2). The median divergence time for the node separating *Dendrobium* / *Earina* from the genus *Cymbidium* was 17.08 Ma and the divergence dates for the genus *Goodyera* radiating from the genus *Cymbidium* was estimated to be 19.3 Ma. The posterior age distribution of the calibrated internal nodes summarized by reading the BEAST generated log files in Tracer v1.5 are shown in Figure. 3.

The phylogeographic tree with divergence rate variation among lineages is depicted in the summarized MCC tree for the homogenous model using a color annotation (Figure 4). All the Asian *Cymbidium* species formed a monophyletic group with three sub clades displaying the three photosynthetic traits viz. weak CAM, strong CAM and C3 dispersed among the taxa. *G. speciosum* from Florida emerged as the outgroup with a weak CAM trait. *C. finlaysonianum* and *C. madidum* with strong CAM and weak CAM traits respectively shared the common ancestor implying a probable transition of weak CAM to strong CAM. The posterior probability density was estimated to calculate the dispersion rate of the Asian *Cymbidium* species with a mean value of 429.36, a 95% HPD lower bound of 149.42 and 95% HPD upper bound of 728.96 (Figure 5). The annotated MCC tree file was imported to SPREAD and the generated KML files were read in Google earth for the visual display of the annotated MCC tree in continuous space (Figure 6).

Discussion

As sequences from multiple species flooded the public domains in late 70s it became apparent that a clock is not always a good model for the process of molecular evolution (Langley and Fitch, 1974). Variation in nucleotide substitution rates, both along a lineage and between different lineages, is known to be ubiquitous and several reasons hypothesize these deviations from the clock-like model of sequence evolution viz.; i) generation time: a lineage with shorter generation time accelerates the clock in the process of accumulation and fixation of new mutations in genetic recombination events (Ohta, 2002), ii) metabolic rates: higher metabolic rates have increased rates of DNA synthesis and higher rates of nucleotide mutations (Gillooly *et al*., 2005), iii) mutation rate: differences in the species-characteristic fidelity of the DNA replication or DNA repair machinery (Ota and Penny, 2003); and iv) effect of effective population size on the rate of fixation of mutations (Ohta, 2002). Hence, it is worth testing the molecular sequence data for rate heterogeneity or the relaxed clock models. The variable rate methods implemented in BEAST uses Bayesian inference and the MCMC procedures to derive the posterior distribution of rates and times. In addition to the auto correlated models like the ones implemented in multidivtime, a range of different, novel models have been implemented, where the rates are drawn from a distribution and these methods employ interesting features such as estimation of the parameter distributions and testing of the correlation of rates between adjacent branches (if > 0, indicates rate inheritance). BEAST also permits the definition of calibration distributions (such as normal, lognormal, exponential or gamma) to model calibration uncertainty instead of simple point estimates or age intervals. For the other, non-calibrated nodes, there is no specific process that describes the prior distribution of divergence times (they are uniform over a range from 0 to very large). BEAST allows the user to simultaneously analyze multiple data sets/partitions with different substitution models, and provides Bayesian credibility intervals.

According to our estimates based on the uncorrelated relaxed molecular clock approach, and incorporating two previously studied internal calibration points, extant Asian *Cymbidium* groups shared a MRCA in the Late Oligocene, about 37.38 Ma. The median age estimates for the crown group subfamilies are consistently younger in our study as compared to the youngest mean ages obtained by Ramirez *et al*. (2007) (Figure 2). A Late Palaeocene radiation (~56 Ma) was proposed by Ramirez *et al.* (2007) for the Orchidaceae, i.e., during the prominent temperature increase of the Early Cenozoic (from 59 Ma to 52 Ma;) era. Our estimations for the genus *Cymbidium* and inclusion of an internal calibration node with *Dendrobium* and *Earina* species which belong to the higher Epindendroides gives a potential explanation for the initial radiation of Orchidaceae during the Eocene as cooler temperatures increased the global heterogeneity of ecosystems, creating new habitats that could conceive adaptive radiation with increasing allopatric speciation (Haffer, 1997).

The crown age estimation of *Cymbidium* indicates a Late Miocene radiation (~11 Ma; Figure 2) based entirely on a secondary calibration point, which should have a direct effect on all internal divergence times. BEAST makes different evolutionary assumptions and have very different algorithms for estimating correct divergence times. One way to assess the autocorrelation assumption is to examine the covariance between parent and child branch in each phylogeny which is calculated by the tool Tracer v1.5 from the MCMC analyses and should be significantly positive when rates are auto correlated, and approach zero when there is no evidence of autocorrelation (refer BEAST manual). For the *Cymbidium* data set, this covariance had a mean of 0.10 and 95% confidence intervals ranging from -0.06 to 0.26 (Figure 3c). The low covariance found in this study does not provide positive evidence for autocorrelation, thus favoring the BEAST analyses projected in this manuscript.

It is worth noting that the single fossil calibration used by Ramirez *et al.* (2007) for tribe Goodyerinae, did influence the age estimates for the *Cymbidium* group in the BEAST analysis which is evident from observing the distribution of ages of the MRCA of tribe Goodyerinae, which falls between 95% HPD lower 1.9Ma and 95%HPD upper 20.2 Ma (Figure 3a).

Phylogeographic Inference

We reconstructed the evolution of photosynthetic characters mapping on a robust phylogenetic hypothesis of Asian *Cymbidium* species by use of the Bayesian phylogeographic analysis (Figure 4). There are different approaches to integrating geography within evolutionary models and a recent approach by Lemey *et al.* (2009) using the BEAST program has gained popularity. BEAST uses Bayesian inferences using MCMC to estimate dispersal throughout the entire evolutionary model, with geographical locations represented as discrete states. The observed geographical data is projected in the tips of the Bayesian phylogenetic tree, and the MCMC method along with a substitution matrix are employed for ancestral states estimation highlighting the migration paths along the branches (Lemey *et al*., 2009). These studies heavily rely on geographical information in their analysis and phylogeography. The amount of geographical information for a molecular sequence in GenBank is underreported in relation to the amount of data that actually exist thereby impeding proper phylogeography analyses. In the present study, the tips section in the BEAUti was fed with approximate sampling time of the Asian *Cymbidium* data in terms of years and months retrieved from the GenBank. There were three sub monophyletic groups with two strong CAM species along with one weak CAM (*C. madidum* sampled from UK) in the molecular phylogenetic tree (Figure 4). There can be two alternative scenarios on the evolution of strong CAM; one is to hypothesize the evolution to strong CAM and subsequent shifts to weak CAM and C3 photosynthesis and alternately to assume parallel evolution to strong CAM at the internode of the common ancestor of *C. bicolor* and *C. finlaysonianum* and at the internode leading to *C. madidum*. The weak CAM species appear in all the three major clades in *Cymbidium* (Figure 4). Weak CAM in *C. dayanum*, *C. devonianum*, and *C. tigrinum* may be interpreted as plesiomorphic while weak CAM in *C. madidum* and *C. chloranthum* is either plesiomorphic or apomorphic under the alternative criteria on the optimization of character evolution. The other apomorphic state, C3 metabolism, was also scattered in the three major clades of the genus (Figure 4). Weak CAM likely shifted to C3 metabolism in the common ancestor of *C. insigne*, *C. erythraeum*, *C. lowianum*, and *C. tracyanum*. Transition from weak CAM to C3 metabolism appeared in the common ancestor of *C. chloranthum*, *C. floribundum*, *C. goeringii*, *C. ensifolium*, *C. kanran*, and *C. lancifolium*. Because the outgroup taxon (*G. speciosum*) studied here is characterized with weak CAM, it would be appropriate to hypothesize that the weak CAM is the ancestral state of photosynthetic characters in *Cymbidium*.

Conclusions

In the last few years molecular dating techniques have greatly improved offering novel opportunities to study the temporal evolution of taxa by fossil data calibration using sequence markers. But there can be concerns regarding the appropriateness and reliability of methodology and other parameters (taxon sampling, fossil calibrations, sequence markers). This study has enabled age estimations for Asian *Cymbidium* (12Ma) for the first time using BEAST by addition of previously analyzed two internal calibration points. Our findings also suggest that the plasticity of photosynthetic character is evolutionarily important. It provides a noteworthy explanation for the fact that *Cymbidium* has been able to occupy ecological niches differing widely in the constitution of environmental factors. On the evolutionary time scale, this plasticity may also provide a genetic reservoir for radiation through changing climatic conditions.

PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

References

1. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2011. GenBank. *Nucleic Acids Research* **39**: D32-7.
2. Bielejec F, Rambaut A, Suchard MA, Lemey P. 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* **27**: 2910-2912.
3. Bremer K, Friis EM, Bremer B. 2004. Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. *Systems Biology* **53**: 496–505.
4. Bremer K. 2000. Early Cretaceous lineages of monocot flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 4707-4711.
5. Britten RJ. 1986. Rates of DNA sequence evolution differ between taxonomic groups. *Science* **231**: 1393–1398.
6. Catresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17**: 540–552.
7. Chaw SM, Chang CC, Chen HL, Li WH. 2004. Dating the monocot–dicot divergence and the origin of core eudicots using whole chloroplast genomes. *Journal of Molecular Evolution* **58**: 424–441.
8. Choi SH, Kim MJ, Lee JS, Ryu KH. 2006. Genetic diversity and phylogenetic relationships among and within species of oriental cymbidiums based on RAPD analysis. *Scientia Horticulturae* 108: 79-85
9. Conran JG, Bannister JM, Lee DE. 2009. Earliest orchid macrofossils: Early Miocene *Dendrobium* and *Earina* (Orchidaceae: Epidendroideae) from New Zealand. *American Journal of Botany* **96**: 466-474
10. Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**: 1969-73

PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

11. Du Puy D, Cribb P. 2007. The Genus *Cymbidium*. Surrey, United Kingdom: Royal Botanic Gardens, Kew Publishing.
12. Du Puy DJ, Cribb P. 1988. The genus *Cymbidium*. Portland: Timber Press
13. Gillooly JF, Allen AP, West GB, Brown JH. 2005. The rate of DNA evolution: Effects of body size and temperature on the molecular clock. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 140–145.
14. Gustafsson AL, Verola CF, Antonelli A. 2010. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). *BMC Evolutionary Biol*ogy **10**:177
15. Haffer J. 1997. Alternative models of vertebrate speciation in Amazonia: An overview. *Biodiversity and Conservation* **6**: 451-476.
16. Janssen T, Bremer K. 2004. The age of major monocot groups inferred from 800+ rbcL sequences. *Botanical Journal of the Linnean Society* **146**: 385-398
17. Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* **9**: 286–298.
18. Langley CH, Fitch W. 1974. An estimation of the constancy of the rate of molecular evolution*. Journal of Molecular Evolution* **3**:161–177.
19. Leigh JW, Lapointe F-J, Lopez P, Bapteste E (2011). Evaluating phylogenetic congruence in the post-genomic era. *Genome Biol Evol* **3**: 571–587
20. Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009. Bayesian phylogeography finds its roots. *PLOS Computational Biology* **5**: e1000520
21. Magallón SA, Sanderson MJ. 2001. Absolute diversification rates in angiosperm clades. *Evolution* **55**: 1762–1780.

22. Motomura H, Yukawa T, Ueno O, Kagawa A. 2008. The occurrence of crassulacean acid metabolism in *Cymbidium* (Orchidaceae) and its ecological and evolutionary implications. *Journal of Plant Research* **121**:163-77
23. Ohta T. 2002. Near-neutrality in evolution of genes and in gene regulation. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 16134–16137.
24. Ota R, Penny D. 2003. Estimating changes in mutational mechanisms of evolution. *Journal of Molecular Evolution* **57**: S233–S240.
25. Rambaut A, Drummond AJ. 2007. TRACER version 1.5. Available: http://beast.bio.ed.ac.uk/Tracer
26. Ramirez SR, Gravendeel B, Singer RB, Marshall CR, Pierce NE. 2007. Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature* **448**: 1042-1045
27. Ronquist F, Huelsenbeck JP (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572-1574
28. Sanderson MJ, Thorne JL, Wikström N, Bremer K. 2004. Molecular evidence on plant divergence times. *American Journal of Botany* **91**: 1656 –1665.
29. Sanderson MJ, Doyle JA. 2001. Sources of error and confidence intervals in estimating the age of Angiosperms from rbcL and 18S rDNA data. *American Journal of Botany* **88**: 1499–1516.
30. Sharma SK, Dkhar J, Kumaria S, Tandon P, Rao SR. 2012. Assessment of phylogenetic inter-relationships in the genus *Cymbidium* (Orchidaceae) based on internal transcribed spacer region of rDNA. *Gene* **13**(1): 10–15
31. Soltis PS, Soltis DE, Savolainen V, Crane PR, Barraclough TG. 2002. Rate heterogeneity among lineages of tracheophytes: Integration of molecular and fossil

data and evidence for molecular living fossils. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 4430 – 4435.

32. Sytsma KJ, Litt A. Zjhra ML, Pires JC, Nepokroeff M, Con Ti E, Walker J, Wilson PG. 2004. Clades, clocks, and continents: historical and biogeographical analysis of Myrtaceae, Vochysiaceae, and relatives in the southern hemisphere. *International Journal of Plant Sciences* **165**: 85–105.
33. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*earch **22**: 4673-4680
34. van den Berg C, Ryan A, Cribb PJ, Chase MW. 2002. Molecular phylogenetics of *Cymbidium* (Orchidaceae: Maxillarieae): Sequence data from internal transcribed spacers (ITS) of nuclear ribosomal DNA and plastid matK. *Lindleyana* **13**: 102–111
35. Wang HZ, Lu JJ, Hu X, Liu JJ. 2011. Genetic variation and cultivar identification in *Cymbidium ensifolium*. *Plant Systematics and Evolution* **13**: 101–110.
36. Wikström N, Savolainen V, Chase MW (2001). Evolution of the Angiosperms: Calibrating the family tree. *Proceedings of the Royal Society B: Biological Sciences* **268**: 2211-2220
37. Wong SC, Hew CS. 1973. Photosynthesis and photorespiration in some thin-leaved orchid species. *Journal of the Singapore National Academy of Science* **3**: 150–157
38. Yoon HS, Hackett JD, Ciniglia C, Pinto G. Bhattacharya, D (2004). A molecular timeline for the origin of photosynthetic Eukaryotes. *Molecular Biology and Evolution* **21**: 809– 818.
39. Zuckerkandl E, Pauling L. 1965. Molecules as documents of evolutionary history. *Journal of Theoretical Biology* **8**: 357-66.

Titles and legends to figures

Figure 1 - Bayesian Phylogram (majority rule consensus tree) inferred from the aligned supermatrix dataset (ITS + matk).

Tree results from a MRBAYES analysis of a 46-species dataset including four combined chloroplast and nuclear gene regions (matK, 5.8S, ITS1 and ITS2). The analysis was run with 10 million generations and a burn-in of 25%. Clade support above 0.49 posterior probability (PP) and 49% bootstrap support (BS) is shown on the tree, and thick branches indicate a support of 1.00 PP/100% BS. The species with known information about their photosynthetic traits are designated as weak CAM, Str CAM and C3. G. speciosum which is a weak CAM is the outgroup.

Figure 2 - Time-calibrated tree focusing on the orchid genus Cymbidium, based on nuclear ribosomal DNA sequences (nrITS), estimated with BEAST.

Thin branches indicate posterior probabilities below 0.9. Numbers at nodes are median ages in million of years (Ma). Node bars indicate the 95% HPD lower and upper bounds in Ma. temporal origin and diversification of the genus Cymbidium, using the Bayesian uncorrelated relaxed molecular clock approach implemented in BEAST with two internal calibration points (15-20 Ma fossil data of an extinct stingless bee as a pollinator for the orchid subtribe Goodyerinae and macromolecular leaf 20 Ma fossil data of genera Dendrobium and Earina). Coloured circles indicate internal calibration point and estimated point of divergence as enumerated in the figure.

PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

Figure 3 - Posterior age distribution of the two fossil-calibrated internal nodes in the orchid phylogeny (outlined in Figure 2) and covariance estimation.

Diagrams show the age of the most common ancestor of these nodes (in millions of years) plotted against frequency of trees, from the combined results analyzed by log combiner and log files read by Tracer v1.5.

Figure 4 - Phylogeographic inference in continuous space model for Asian Cymbidium spp from North east India and China

MCC phylogeny with branches colored according to the most probable posterior location of their child nodes; superimposed under the phylogeny lies the inferred demographic history of Asian cymbidium. The axes represnts sampling time of the molecular data from the respectives Cymbidium species as per the information from GenBank records.

Figure 5 - Posterior age distribution plots for the Asian Cymbidium phylogeographic inference showing dispersion rates

Diagrams show the posterior probability density estimating the dispersion rate of the Asian Cymbidum species with a mean value of 429.36, a 95% HPD lower bound of 149.42 and 95% HPD upper bound of 728.96

Figure 6 - Screenshot of SPREAD showing Asian Cmbidium phylogeography in continuous space by incorporating MCC tree files from BEAST and Tree Annotator into SPREAD for virtual display in Google Earth

Visualization of phylogeographic history of Asian Cymbidium spp under a continuous diffusion model. Generation of KML output enables interactive exploration in the time dimension as well in freely available virtual globe software, google earth.

Table 1. ITS and matK accessions for *Cymbidium* species

Sl. No.	**Species**	**Accession number for ITS**	**Accession number for matK**
1.	*Cymbidium aloifolium*	AJ300269.1	AF470485
2.	*Cymbidium goeringii*	AJ300289.1	JX202671.1
3.	*Cymbidium forrestii*	AJ300287.1	matK sequences not available in
4.	*Cymbidium formosanum*	AJ300283.1	matK sequences not available in
5.	*Cymbidium kanran*	AJ300281.1	JX202672.1
6.	*Cymbidium faberi*	AJ300275.1	matK sequences not available in
7.	*Cymbidium ensifolium*	AJ300273.1	AF470464.1
8.	*Cymbidium lowianum*	AJ300271.1	AF470477.1
9.	*Cymbidium aspidistrifolium*	AJ300286.1	matK sequences not available in
10.	*Cymbidium rubrigemmum*	AJ300284.1	matK sequences not available in
11.	*Cymbidium gyokuchin*	AJ300278.1	matK sequences not available in
12.	*Cymbidium sinense*	AJ300272.1	HM137056.1
13.	*Cymbidium insigne*	AJ300270.1	AF470476.1
14.	*Cymbidium cochleare*	JN114487.1	matK sequences not available in
15.	*Cymbidium devonianum*	JN114488.1	HM137053.1
16.	*Cymbidium giganteum*	JF729013.1	AY368402.1
17.	*Cymbidium eburneum*	JF729012.1	AF470470.1
18.	*Cymbidium hookerianum*	JF729010.1	AF470471.1
19.	*Cymbidium mastersii*	JF729006.1	AF470481.1
20.	*Cymbidium lancifolium*	AF284724.1	AF470475.1
21.	*Cymbidium defoliatum*	AF284718.1	matK

PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

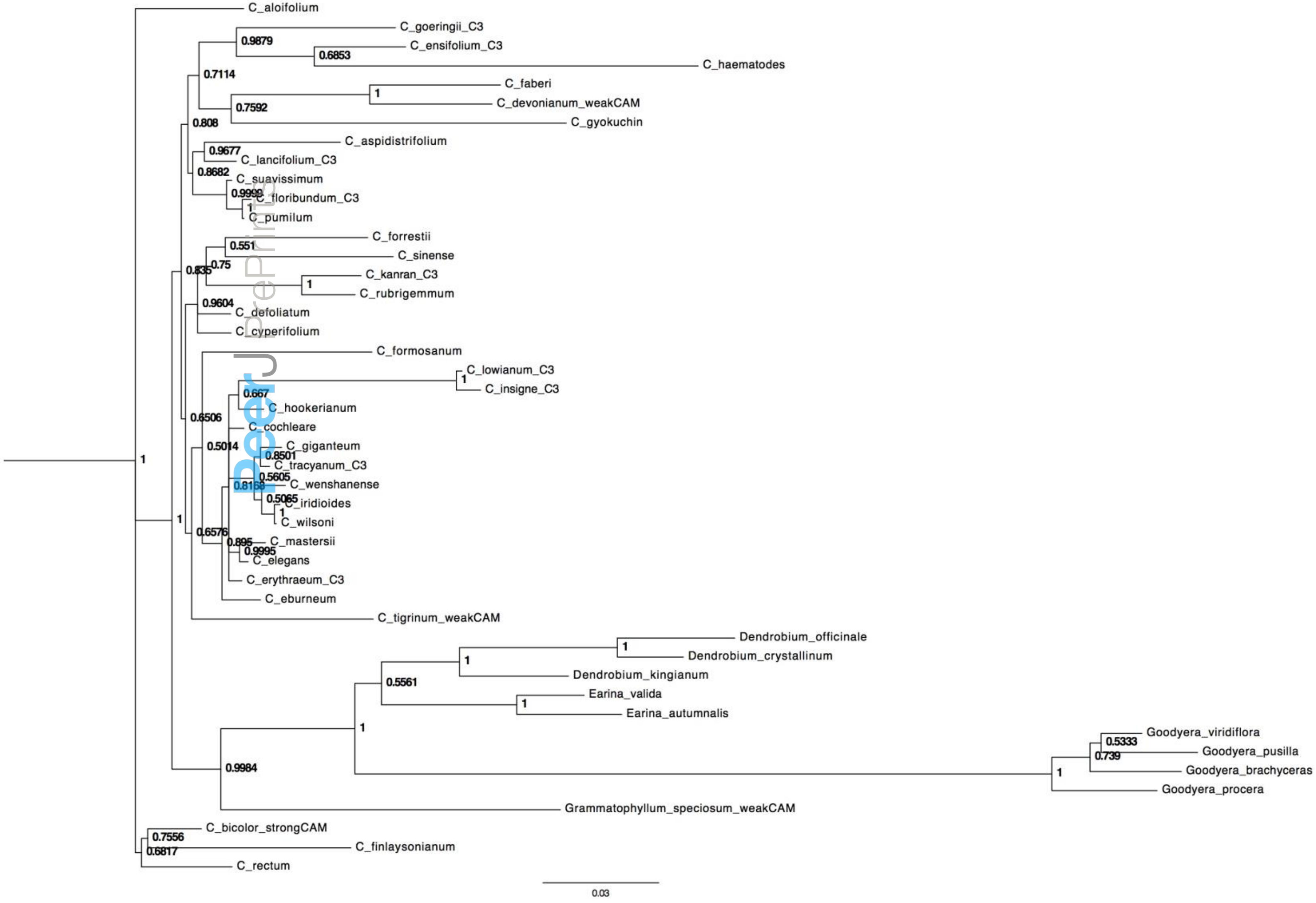
			sequences not available in genBank
22.	*Cymbidium elegans*	AF284712.1	AF470478.1
23.	*Cymbidium wenshanense*	AF284708.1	matK sequences not available in
24.	*Cymbidium iridioides*	AF284702.1	AF470468.1
25.	*Cymbidium suavissimum*	AF284700.1	AF470472.1
26.	*Cymbidium floribundum*	AF284698.1	AF470467.1
27.	*Cymbidium bicolor*	AF284696.1	GU990531.1
28.	*Cymbidium cyperifolium*	AF284719.1	matK sequences not available in GenBank
29.	*Cymbidium tigrinum*	AF284713.1	HM137054.1
30.	*Cymbidium wilsoni*	AF284705.1	matK sequences not available in GenBank
31.	*Cymbidium erythraeum*	AF284703.1	AF470469.1
32.	*Cymbidium tracyanum*	AF284701.1	AF470484.1
33.	*Cymbidium pumilum*	AF284699.1	matK Not available in GenBank
34.	*Dendrobium officinale*	HQ114245.1	GU181388.1
35.	*Dendrobium crystallinum*	HQ114243.1	AF445447.1
36.	*Dendrobium kingianum*	EU430386.1	AF263651.1
37.	*Goodyera viridiflora*	JN166067.1	JN166027.1
38.	*Goodyera procera*	JN114517.1	JN004450.1
39.	*Goodyera brachyceras*	AM778169.1	AM902104.1
40.	*Goodyera pusilla*	ITS2 sequences not available in GenBank	JN166026.1
41.	*Earina valida*	AF521077.1	AY121741.1
42.	*Earina autumnalis*	AF260149.1	EF079336.1
43.	*Cymbidium rectum*	ITS2 sequences not available in GenBank	AF470463.1
44.	*Cymbidium finlaysonianum*	-do-	HM137048.1
45.	*Cymbidium haematodes*	-do-	HM137063.1

Table 2. priors for model parameters and statistics settings for the concatenated datset to be run BEAST.

Parameter	**Prior**	**Bound**	**Description**
tmrca(Cymbidium)	Using Tree Prior	n/a	tmrca statistic for taxon set untitled1 on tree cymbTree
tmrca(Dendrobium_Earina)	Uniform [20, 120]	n/a	tmrca statistic for taxon set untitled2 on tree cymbTree
tmrca(Goodyera)	Exponential [10]	n/a	tmrca statistic for taxon set untitled3 on tree cymbTree
Matrix in file "cymb_ITS1_2.nexus".ac	Gamma [0.05, 10], initial=1	[0, ∞]	GTR A-C substitution parameter
Matrix in file "cymb_ITS1_2.nexus".ag	* Gamma [0.05, 20], initial=1	[0, ∞]	GTR A-G substitution parameter
Matrix in file "cymb_ITS1_2.nexus".at	* Gamma [0.05, 10], initial=1	[0, ∞]	GTR A-T substitution parameter
Matrix in file "cymb_ITS1_2.nexus".cg	* Gamma [0.05, 10], initial=1	[0, ∞]	GTR C-G substitution parameter
Matrix in file "cymb_ITS1_2.nexus".gt	* Gamma [0.05, 10], initial=1	[0, ∞]	GTR G-T substitution parameter
Matrix in file "cymb_ITS1_2.nexus".frequencies	Uniform [0, 1], initial=0.25	[0, 1]	base frequencies
Matrix in file "cymb_ITS1_2.nexus".alpha	* Exponential [0.5], initial=0.5	[0, ∞]	gamma shape parameter
Matrix in file "cymb_ITS1_2.nexus".pInv	* Uniform [0, 1], initial=0.5	[0, 1]	proportion of invariant sites parameter
Matrix in file "cymbidium_matk.nexus".ac	* Gamma [0.05, 10], initial=1	[0, ∞]	GTR A-C substitution parameter
Matrix in file "cymbidium_matk.nexus".ag	* Gamma [0.05, 20], initial=1	[0, ∞]	GTR A-G substitution parameter
Matrix in file "cymbidium_matk.nexus".at	* Gamma [0.05, 10], initial=1	[0, ∞]	GTR A-T substitution parameter
Matrix in file	* Gamma [0.05, 10],	[0, ∞]	GTR C-G

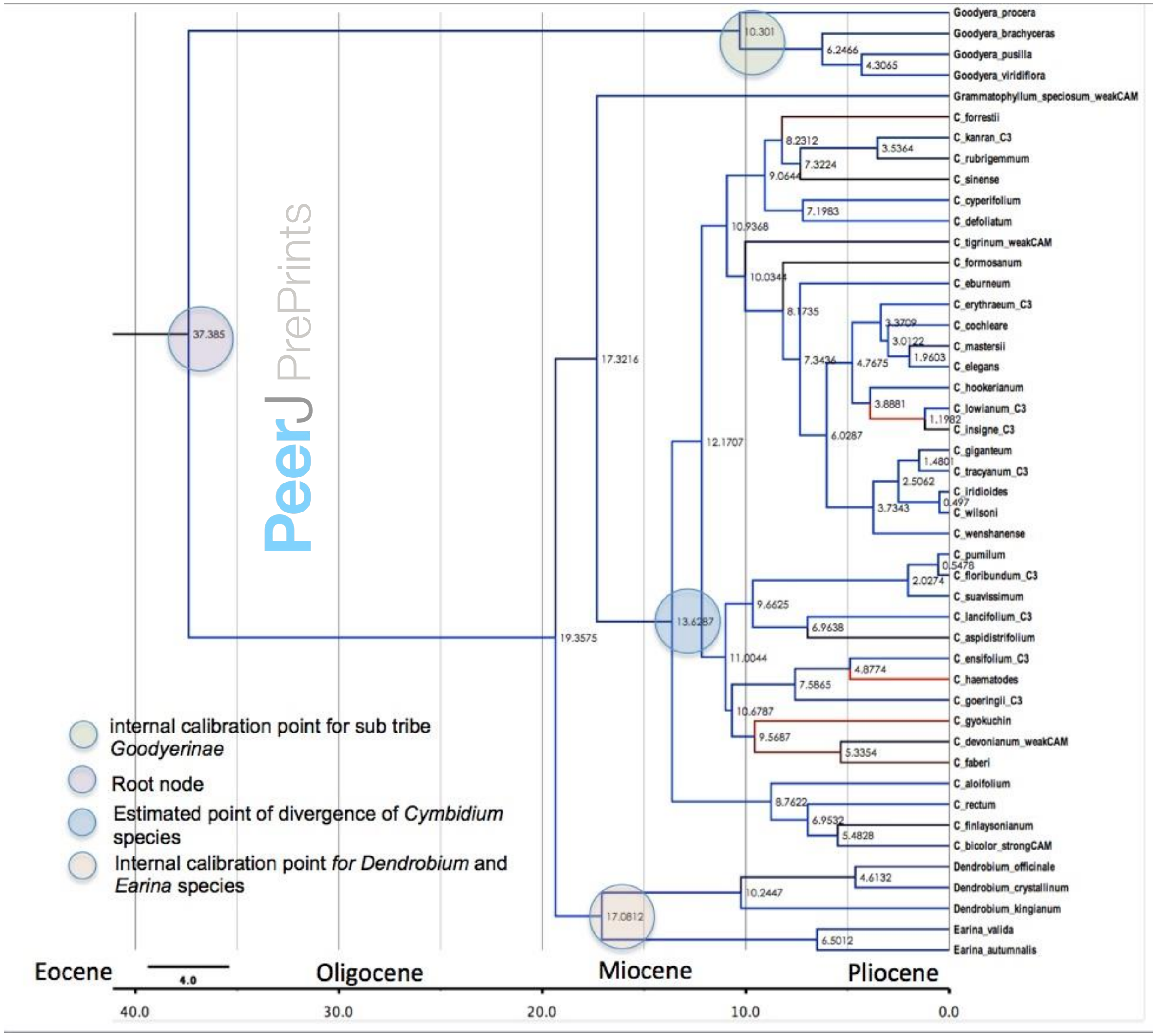
"cymbidium_matk.nexus".cg	initial=1		substitution parameter
Matrix in file "cymbidium_matk.nexus".gt	* Gamma [0.05, 10], initial=1	[0, ∞]	GTR G-T substitution parameter
Matrix in file "cymbidium_matk.nexus".frequencies	* Uniform [0, 1], initial=0.25	[0, 1]	base frequencies
Matrix in file "cymbidium_matk.nexus".alpha	* Exponential [0.5], initial=0.5	[0, ∞]	gamma shape parameter
Matrix in file "cymbidium_matk.nexus".pInv	* Uniform [0, 1], initial=0.5	[0, 1]	proportion of invariant sites parameter
ucld.stdev	Exponential [10], initial=1	[0, ∞]	uncorrelated lognormal relaxed clock stdev
ucld.mean	Uniform [1E-6, 10], initial=1	[1E-6, 10]	uncorrelated lognormal relaxed clock mean
treeModel.rootHeight	Normal [39, 5.5], initial=0.034	[0, ∞]	root height of the tree
yule.birthRate	Uniform [0, 1E100], initial=0.4	[0, 1E100]	Yule speciation process birth rate
meanRate	* Indirectly Specified Through Other Parameter	n/a	The mean rate of evolution over the whole tree
covariance	* Indirectly Specified Through Other Parameter	n/a	The covariance in rates of evolution on each lineage with their ancestral lineages
coefficientOfVariation	* Indirectly Specified Through Other Parameter	n/a	The variation in rate of evolution over the whole tree

Figure 1



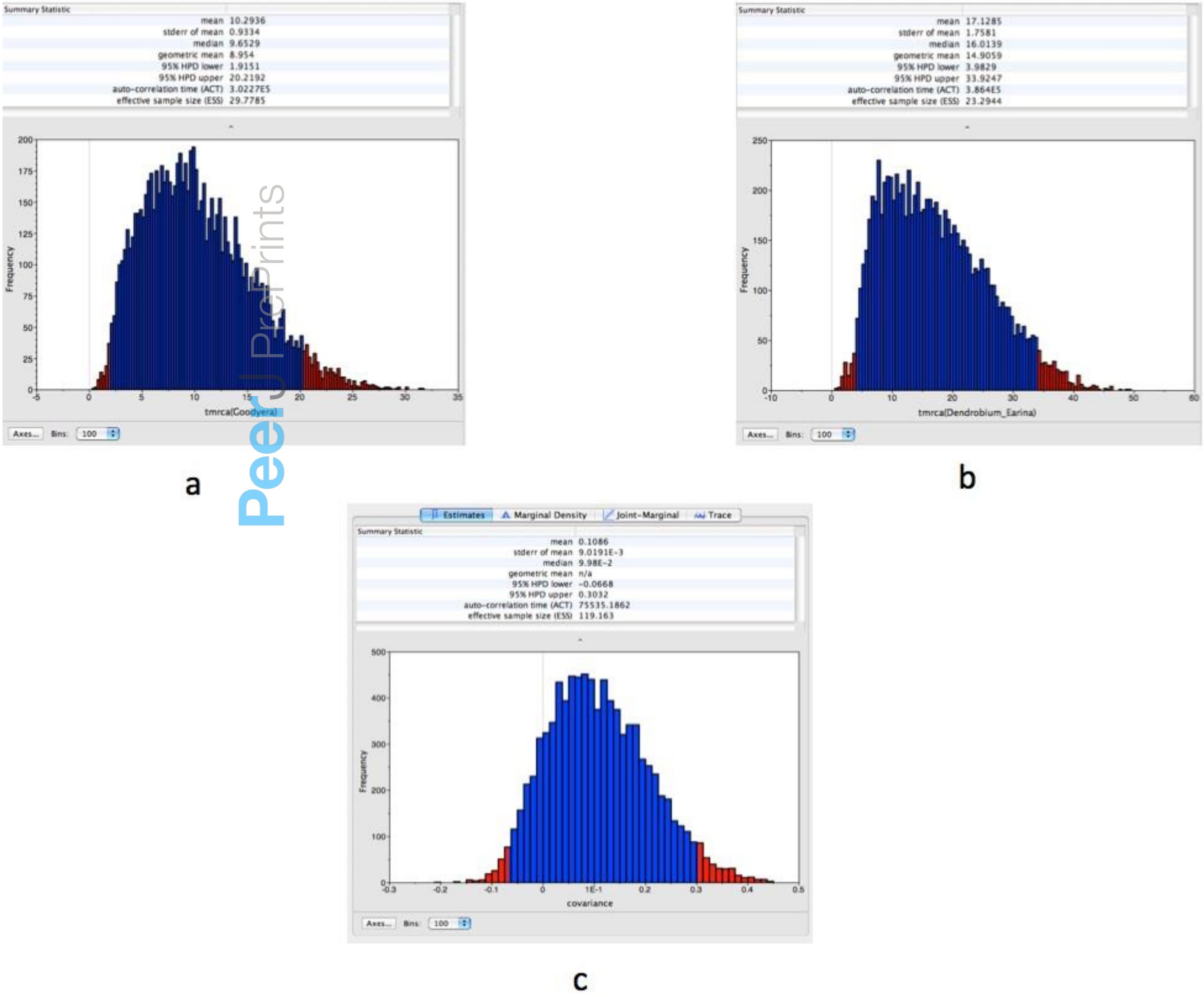
PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

Figure 2.



PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

Figure 3.



```
Summary Statistic
                          mean  10.2936
                stderr of mean  0.9334
                        median  9.6529
                geometric mean  8.954
                 95% HPD lower  1.9151
                 95% HPD upper  20.2192
   auto-correlation time (ACT)  3.0227E5
  effective sample size (ESS)  29.7785
```

a

```
Summary Statistic
                          mean  17.1285
                stderr of mean  1.7581
                        median  16.0139
                geometric mean  14.9059
                 95% HPD lower  3.9829
                 95% HPD upper  33.9247
   auto-correlation time (ACT)  3.864E5
  effective sample size (ESS)  23.2944
```

b

```
Summary Statistic
                          mean  0.1086
                stderr of mean  9.0191E-3
                        median  9.98E-2
                geometric mean  n/a
                 95% HPD lower  -0.0668
                 95% HPD upper  0.3032
   auto-correlation time (ACT)  75535.1862
  effective sample size (ESS)  119.163
```

c

PeerJ PrePrints | https://peerj.com/preprints/94v1/ | v1 received: 9 Nov 2013, published: 9 Nov 2013, doi: 10.7287/peerj.preprints.94v1

Figure 4.

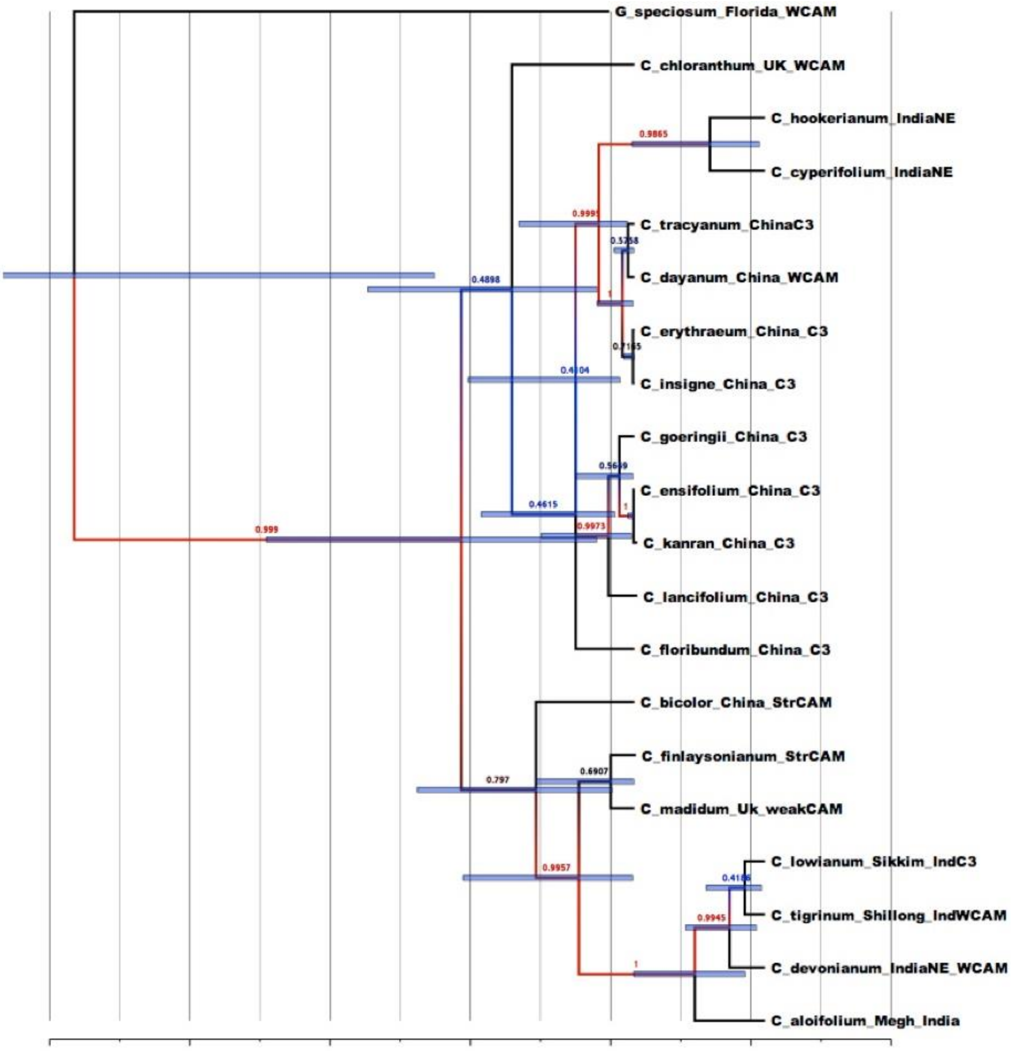
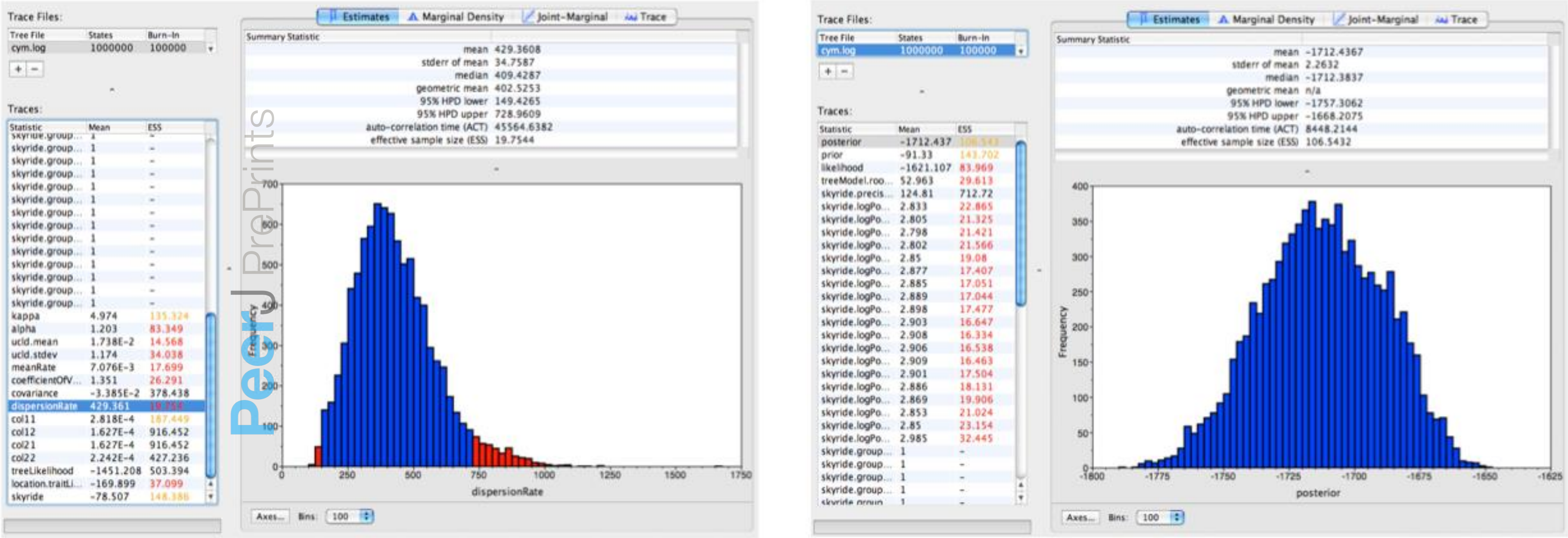

G_speciosum_Florida_WCAM
C_chloranthum_UK_WCAM
0.9865
C_hookerianum_IndiaNE
C_cyperifolium_IndiaNE
0.9995
C_tracyanum_ChinaC3
0.5758
0.4898
1
C_dayanum_China_WCAM
C_erythraeum_China_C3
0.7165
0.4104
C_insigne_China_C3
C_goeringii_China_C3
0.5649
C_ensifolium_China_C3
0.4615
1
0.999
0.9973
C_kanran_China_C3
C_lancifolium_China_C3
C_floribundum_China_C3
C_bicolor_China_StrCAM
C_finlaysonianum_StrCAM
0.797
0.6907
C_madidum_Uk_weakCAM
0.9957
C_lowianum_Sikkim_IndC3
0.4186
C_tigrinum_Shillong_IndWCAM
0.9945
1
C_devonianum_IndiaNE_WCAM
C_aloifolium_Megh_India

Figure 5.



Trace Files:

Tree File	States	Burn-In
cym.log	1000000	100000

Traces:

Statistic	Mean	ESS
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
kappa	4.974	135.324
alpha	1.203	83.349
ucld.mean	1.738E-2	14.568
ucld.stdev	1.174	34.038
meanRate	7.076E-3	17.699
coefficientOfV...	1.351	26.291
covariance	-3.385E-2	378.438
dispersionRate	429.361	19.754
col11	2.818E-4	187.449
col12	1.627E-4	916.452
col21	1.627E-4	916.452
col22	2.242E-4	427.236
treeLikelihood	-1451.208	503.394
location.traitLi...	-169.899	37.099
skyride	-78.507	148.386

Summary Statistic	
mean	429.3608
stderr of mean	34.7587
median	409.4287
geometric mean	402.5253
95% HPD lower	149.4265
95% HPD upper	728.9609
auto-correlation time (ACT)	45564.6382
effective sample size (ESS)	19.7544

a

Trace Files:

Tree File	States	Burn-In
cym.log	1000000	100000

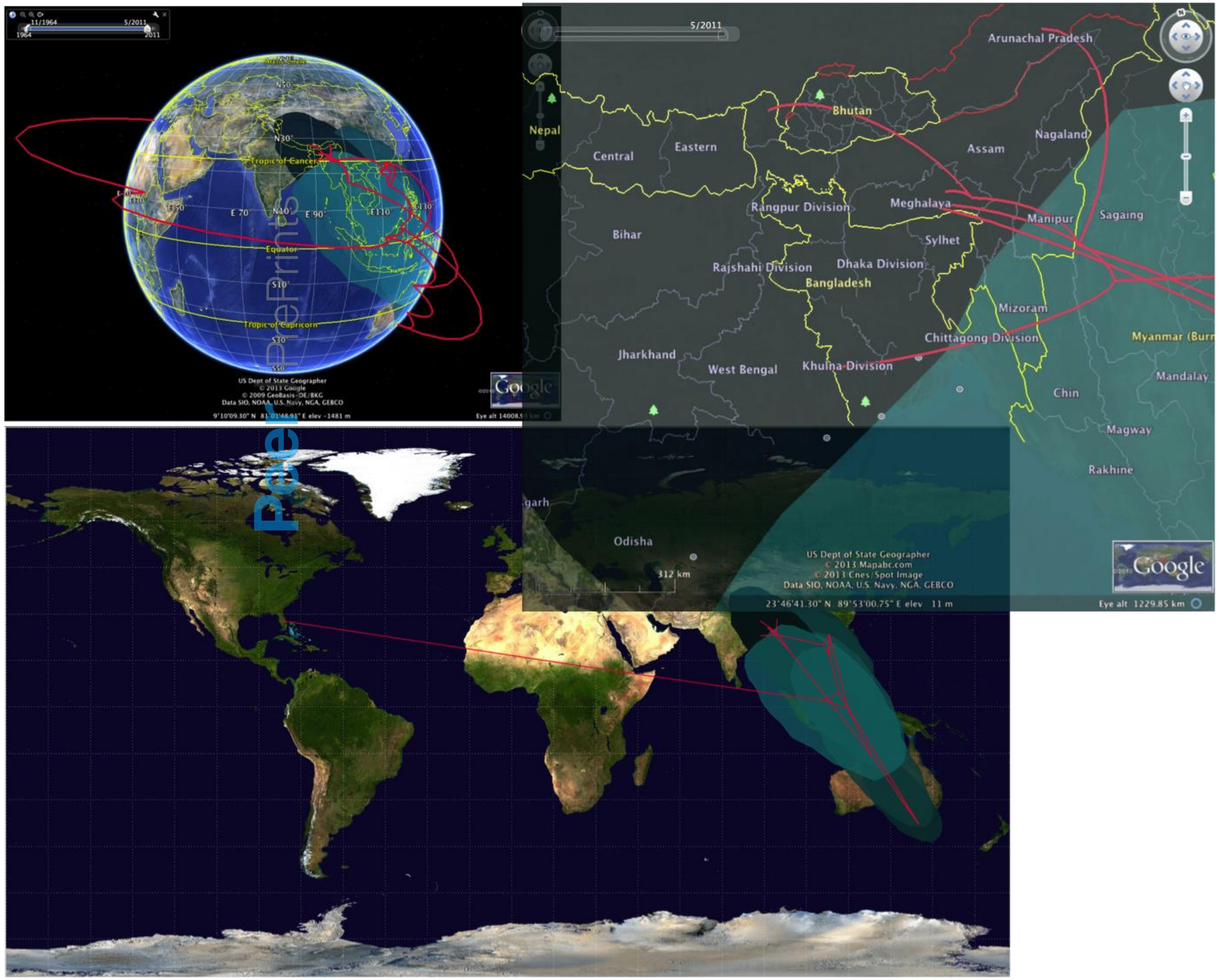
Traces:

Statistic	Mean	ESS
posterior	-1712.437	106.543
prior	-91.33	143.702
likelihood	-1621.107	83.969
treeModel.roo...	52.963	29.613
skyride.precis...	124.81	712.72
skyride.logPo...	2.833	22.865
skyride.logPo...	2.805	21.325
skyride.logPo...	2.798	21.421
skyride.logPo...	2.802	21.566
skyride.logPo...	2.85	19.08
skyride.logPo...	2.877	17.407
skyride.logPo...	2.885	17.051
skyride.logPo...	2.889	17.044
skyride.logPo...	2.898	17.477
skyride.logPo...	2.903	16.647
skyride.logPo...	2.908	16.334
skyride.logPo...	2.906	16.538
skyride.logPo...	2.909	16.463
skyride.logPo...	2.901	17.504
skyride.logPo...	2.886	18.131
skyride.logPo...	2.869	19.906
skyride.logPo...	2.853	21.024
skyride.logPo...	2.85	23.154
skyride.logPo...	2.985	32.445
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group...	1	-
skyride.group	1	-

Summary Statistic	
mean	-1712.4367
stderr of mean	2.2632
median	-1712.3837
geometric mean	n/a
95% HPD lower	-1757.3062
95% HPD upper	-1668.2075
auto-correlation time (ACT)	8448.2144
effective sample size (ESS)	106.5432

b

Figure 6.


11/1964
5/2011
1964
2011
Arctic Circle
N50
N30
Tropic of Cancer
E 50
E 70
N10
E 90
E 110
E 130
Equator
S10
Tropic of Capricorn
S30
S50
US Dept of State Geographer
© 2013 Google
© 2009 GeoBasis-DE/BKG
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
9°10'09.30" N 81°01'48.91" E elev -1481 m
Eye alt 14008.93 km
5/2011
Arunachal Pradesh
Nepal
Bhutan
Central
Eastern
Assam
Nagaland
Meghalaya
Rangpur Division
Manipur
Sagaing
Bihar
Sylhet
Rajshahi Division
Dhaka Division
Bangladesh
Mizoram
Chittagong Division
Myanmar (Burr
Jharkhand
West Bengal
Khulna Division
Mandalay
Chin
Magway
Rakhine
garh
Odisha
312 km
US Dept of State Geographer
© 2013 Mapabc.com
© 2013 Cnes/Spot Image
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
23°46'41.30" N 89°53'00.75" E elev 11 m
Eye alt 1229.85 km
Google

From: InCoB2013 <incob2013@easychair.org>
Subject: InCoB2013 notification for paper 82 (BMC Bioinformatics)
Date: 22 June 2013 10:59:19 AM GMT+05:30
To: Pramod Tandon <profptandon@gmail.com>

Dear Pramod Tandon:

We have evaluated all the comments received and concluded that your paper

Submission: 82
Title: Age estimation for the the genus Cymbidium (Orchidaceae: Epidendroideae) with implementation of fossil data calibration using molecular markers (ITS2 & matK) with a Bayesian relaxed clock implying phylogeographic inference in continuous space.

cannot be considered for publication in BMC Bioinformatics InCoB2013 Supplement Issue. The decision after careful consideration of the reviewers' comments is "reject after review". The decision does not reflect in anyway doubt about the scientific methodology and merit of the work. We recommend that you consider submission to a journal specialized in evolutionary biology. Since BMC Bioinformatics has strictly limited the number of articles for this supplement to ten we are unable to accomodate your submission. Only top scoring submissions that fully align with scope of the journal have been accepted.

If you wish to present your work as SHORT TALK within InCoB regular sessions we are delighted to accept an reformatted abstract using the poster template template (http://ictbi.imedcn.org/incob2013/download/Poster_Abstract_Template.doc).

InCoB2013, in cooperation with Faculty of 1000 (http://f1000.com), invites all members to deposit any of their poster/oral presentation(s) slides into the established open access poster repository, F1000 Posters (http://f1000.com/posters) to allow those who could not make the meeting have the opportunity to see your novel work. F1000's expert Faculty of 10,000 members will then view these submissions to identify those they wish to select for positive evaluation in the award winning F1000 service. To deposit your work, simply go to http://f1000.com/posters/depositor and upload your file to maximize the value of your conference presentations.

Kindly note, if you accept our offer you are requested to submit your abstract by June 25.

At least one author of the poster abstract must register for InCoB2013 for a poster presentation of their work, by July 15, 2013. Registration fee payment must be received in full before your abstract can be included in the on-line and printed version of the program and poster abstracts. Similarly, invitation letter for visa application purposes will be only issued if registration fee payment has been received in full.

If you choose not submit an abstract we delete your submission from the Easychair system on June 26.

We look forward to your participation and presentation at InCoB2013.

Christian Schoenbach, Shoba Ranganathan and Bairong Shen
InCoB2013 Publication Co-chairs

---------------------- REVIEW 1 ---------------------
PAPER: 82
TITLE: Age estimation for the the genus Cymbidium (Orchidaceae: Epidendroideae) with implementation of fossil data calibration using molecular markers (ITS2 & matK) with a Bayesian relaxed clock implying phylogeographic inference in continuous space.
AUTHORS: Devendra Biswal, Jean Valrie and Pramod Tandon

OVERALL EVALUATION: 1 (weak accept)
REVIEWER'S CONFIDENCE: 4 (high)

----------- REVIEW -----------
This manuscript presents a study of age estimation for the genus Cymbidum using BEAST coupled with two internal calibration points. The authors have conducted a systematic bioinformatics investigation including data filtering and collection, sequence analysis, Bayesian analysis, dating analysis and locatin analysis. The method is reasonable and the reulst is acceptable.

However, it’s lack of comparative evaluation between other similar work and the presented method in this manuscript. I would suggest the authors could conduct the conparison or at least discuss it.

---------------------- REVIEW 2 ---------------------
PAPER: 82
TITLE: Age estimation for the the genus Cymbidium (Orchidaceae: Epidendroideae) with implementation of fossil data calibration using molecular markers (ITS2 & matK) with a Bayesian relaxed clock implying phylogeographic inference in continuous space.
AUTHORS: Devendra Biswal, Jean Valrie and Pramod Tandon

OVERALL EVALUATION: -1 (weak reject)
REVIEWER'S CONFIDENCE: 3 (medium)

----------- REVIEW -----------
Novelty/originality

'Yes' in terms of results, but 'No' in terms of methodolgy (merely application of routine methods to existing data)

Importance to field
weak (bioinformatics); strong (evolutionary biology)

Appropriatness for this journal

better suited for journal molecular evolution

Sound methodology
yes

Support of discussion/conclusions by results
yes

References to prior work
yes

Length, organization and clarity (language)
No. This manuscript should be edited for language and style

Quality of display items
acceptable

Response to reviewers' comments on Paper 81 submitted to BMC Genomics through INCOB 2013 easychair

Dear Sir

Re: Submission Paper 81

Title: Age estimation for the genus *Cymbidium* (Orchidaceae: Epidendroideae) with implementation of fossil data calibration using molecular markers (ITS2 & matK) with a Bayesian relaxed clock implying phylogeographic inference in continuous space

We would like to thank the reviewers for their time and their valuable comments. The reviewers' comments were highly encouraging and enabled us get accepted in PeerJ. In the following lines are our point-by-point responses to the comments of the reviewers.

Reviewer 1

This manuscript presents a study of age estimation for the genus Cymbidum using BEAST coupled with two internal calibration points. The authors have conducted a systematic bioinformatics investigation including data filtering and collection, sequence analysis, Bayesian analysis, dating analysis and location analysis. The method is reasonable and the result is acceptable.

However, it's lack of comparative evaluation between other similar work and the presented method in this manuscript. I would suggest the authors could conduct the conparison or at least discuss it.

Response: The paper is slightly modified with inclusion of comparisons between other similar work and our findings.

Reviewer 2: Favorable comments

We hope that the revisions in the manuscript and our accompanying responses will be sufficient to make our manuscript suitable for publication in PeerJ which was already presented in INCOB 2013 in oral sessions.

Sincerely,

Pramod Tandon

(Corresponding author)